# A Nonlinear Predictive Control Approach for Urban Drainage Networks Using Data-Driven Models and Moving Horizon Estimation

Balla, Krisztian Mark; Schou, Christian; Bendtsen, Jan Dimon; Ocampo-Martinez, Carlos; Kallesøe, Carsten

*Published in:*
I E E E Transactions on Control Systems Technology

*DOI (link to publication from Publisher):*
10.1109/TCST.2021.3137712

*Publication date:*
2022

*Document Version*
Accepted author manuscript, peer reviewed version

# A Nonlinear Predictive Control Approach for Urban Drainage Networks Using Data-Driven Models and Moving Horizon Estimation

Krisztian Mark Balla[a,b], Christian Schou[a], Jan Dimon Bendtsen[b], Carlos Ocampo-Martinez[c] and Carsten Skovmose Kallesøe[a,b]

[a]Grundfos Holding A/S, Poul Due Jensens Vej 7, DK-8850, Bjerringbro, Denmark
[b]Section for Automation and Control, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark
[c]Automatic Control Department, Institut de Robotica i Informatica Industrial (CSIC-UPC), Universitat Politècnica de Catalunya, Llorens i Artigas 4-6, Planta 2, 08028, Barcelona, Spain

***Abstract*—*Real-time control of urban drainage networks is a complex task where transport flows are non-pressurized and therefore impose flow-dependent time delays in the system. Unfortunately, the installation of flow sensors is economically out of reach at most utilities, although knowing volumes and flows are essential to optimize system operation. In this article, we formulate joint parameter and state estimation based on level sensors deployed inside manholes and basins in the network. We describe the flow dynamics on the main pipelines by the level variations inside manholes, characterized by a system of coupled partial differential equations. These dynamics are approximated with kinematic waves where the network model is established with the water levels being the system states. Moving horizon estimation is developed where the states and parameters are obtained via the levels and estimated flow data, utilizing the topological layout of the network. The obtained model complexity is kept within practically achievable limits, suitable for nonlinear predictive control. The effectiveness of the control and estimation method is demonstrated on a high-fidelity model of a drainage network, acting as virtual reality. We use real rain and wastewater flow data and test the controller against the uncertainty in the disturbance forecasts.*

***Keywords*—*Receding horizon control; Transport delay; Partial differential equation; Urban drainage network***



Published in *IEEE Transactions on Control Systems Technology.*

DOI: https://doi.org/10.1109/TCST.2021.3137712

# Contents


1 Introduction . . . . . 1
1.1 Nomenclature . . . . . 3
2 Drainage System overview . . . . . 5
3 System model . . . . . 5
3.1 Physical transport model . . . . . 5
3.2 Reduced, data-driven transport model . . . . . 7
3.3 Disturbance model . . . . . 9
3.4 Storage model . . . . . 10
3.5 Network description . . . . . 11
3.6 Discrete network model . . . . . 12
4 Moving Horizon Estimation . . . . . 12
4.1 Parameter estimation . . . . . 13
4.2 State estimation . . . . . 15
5 Control design . . . . . 16
5.1 NMPC problem . . . . . 16
5.2 Objectives . . . . . 18
6 Numerical results . . . . . 20
6.1 Baseline controller . . . . . 20
6.2 Case study . . . . . 22
6.3 Simulation environment . . . . . 23
6.4 Identification results . . . . . 23
6.5 Control results . . . . . 25
7 Discussion . . . . . 30
8 Conclusions and future work . . . . . 30
References . . . . . 32

# 1 Introduction

Open - channel hydraulic systems are large-scale networks where water is transported with a free surface in pipes or conduits [Schütze et al., 2002]. In this work, we focus on Urban Drainage Networks (UDNs), more specifically, on systems where rain and wastewater are combined, and pumped in open channels. Pumped UDNs are typical in areas where the gravitation of water is limited due to the flatness of the landscape [García et al., 2015]. Moreover, combined sewers carry both domestic and stormwater towards treatment plants, where the sewage is treated before being released to the environment [Butler and Davies, 2006]. Combined sewers are present in many large cities and they are often overloaded due to the under-dimensioned capacity of the infrastructure induced by fast urbanization and the growing number of end-users [Lund et al., 2018]. Besides, UDNs are increasingly being pushed to their limits due to changing weather conditions, resulting in more frequent Combined Sewer Overflows (CSOs) [Lund et al., 2018]. The changing conditions challenge flow prediction and raise the question of how to handle the increased load on these systems.

In the last few decades, several Real-Time Control (RTC) techniques have been developed for UDN applications. These techniques typically exploit the available sensor measurements, rain forecasts, and the available physical description of the network. Many of the applied methods for evaluating the network capacities and solving optimization problems are typically predictive model-based control techniques [Mays, 2001; Ocampo-Martinez et al., 2013; Ocampo-Martinez, 2010]. However, transport flows in open-channel hydraulics are governed by a set of Partial Differential Equations (PDEs), too complex to identify with data and often infeasible to adapt to RTC applications in large-scale problems. Several methods in the literature typically propose the use of reduced PDE-based models in Model Predictive Control (MPC). These methods rely on the physical properties available (e.g., pipe dimensions, friction, and slope parameters) for model calibration with HiFi (High Fidelity) model simulators [Xu et al., 2011], [Xu et al., 2012]. Simulating gravity-driven flow with full PDE-based models in large-scale UDNs requires either a HiFi simulation environment or the placement of several flow and level sensors along the pipelines, meaning prohibitively expensive installation and maintenance costs.

PDEs linearized around an operating point have been used in UDN applications, where transfer functions [Dalmas et al., 2017], [Litrico and Fromion, 2006] and state-space models [Zou et al., 2015] have been developed. Due to the complexity of PDE-based control, conceptual models are also used in the state-of-art, for instance, [Balla et al., 2020b] used algebraic models with a single delay parameter, while [Ocampo-Martinez, 2010], [Gelormino and Ricker, 1994], [Mollerup et al., 2016], [Ocampo-Martinez et al., 2013], [Joseph-Duran et al., 2015] used a dynamic control model where the available capacity of pipes and tanks have been collectively modelled as virtual buffers. However, linearized and conceptual internal models do not allow flow-dependent time

delays, conceptualize the physically measurable levels and flows, furthermore restrict the flow deviation from steady-state solutions. Data-driven modeling has been reported in [Balla et al., 2020a] and in [Troutman et al., 2017], where grey-box and black-box identification have been used, respectively.

In this article, we propose a PDE-based modeling framework, where the system of PDEs is approximated to obtain a simple representation of the network, preserving the main system dynamics for control. We report on the modularity of the framework by arguing that using the network topology and water level sensors, a model suitable for control is obtained. Opposed to the current state-of-art, the proposed modeling framework captures the inflows to the UDN through water level measurements. In this way, we disregard the use of HiFi simulation models for model calibration.

Moreover, a new Nonlinear Model Predictive Control (NMPC) approach is proposed, based on a data-driven model, reduced from PDEs. In our approach, a Moving Horizon Estimation (MHE) method is used for constrained parameter and state estimation governing the PDEs, which we spatially discretize to Ordinary Differential Equations (ODEs). Time periodicity conditions are imposed on disturbance inflows, generated by household activity, to incorporate additional structure in the model dynamics used for predictions in the NMPC. The proposed control architecture is shown in Figure 1. The Moving Horizon Parameter Estimation (MHPE) along with the Moving Horizon State Estimation (MHSE) is carried out using easy-accessible level sensors distributed and placed inside manholes along the main sewer lines. Besides, we utilize flow estimation techniques which allow us to use pumped inlet and gravitated discharge flows, further detailed in [Kallesøe and Knudsen, 2016]. By using MHPE with NMPC, the system can re-identify the slowly changing pipe dynamics due to accumulated sludge in the bottom of sewer pipes. Besides, the NMPC can adapt to varying flow conditions imposed by the changing rain infiltration due to seasonality. The MHPE and MHSE problems, similarly to [Joseph-Duran et al., 2015], [Joseph-Duran et al., 2014], are both formulated as nonlinear least-squares problems, subject to state and parameter constraints, further detailed in Section 4. As shown in Figure 1, the NMPC is utilized as a global controller, solving a multi-criteria optimization problem and thereby providing references to the pumps at the local pumping stations. The proposed control and estimation methods are demonstrated on a HiFi network, simulated in the Mike Urban (MU)[1] simulation software where we use the catchment dynamics and the MU runoff engine for generating rain-runoff appearing as the load on the network. Finding the rain-runoff based on rain intensity forecasts by radars and numerical weather predictions is an active field of research, which has been extensively studied in [Ma et al., 2018], [Chang et al., 2001] and its effect of uncertainty on UDNs in [Löwe et al., 2016], [Löwe et al., 2014]. Moreover, several works in the literature report on how to handle rain forecast uncertainty, e.g.,

[1]MIKE Urban is a standard hydraulic simulation and planning tool, used as a planning tool by many operators at water utilities. The MU simulation environment solves the full dynamic PDEs for open-channel flow [MIKE powered by DHI, 2017].

in UDNs in [Balla et al., 2020b] and in river applications in [Tian et al., 2017]. In this work, historical events of rain and wastewater are utilized in terms of real measurements, representing the imperfect weather forecasts.

The proposed data-driven method using the reduced network model and MHE has two clear benefits:

- First, it is a data-driven method that does not require heavy computation and difficult calibration procedures opposed to HiFi models, used at many utilities.
- Second, it is only required to collect data from periods under normal operational behavior, opposed to conventional data-driven methods where historical data is required for all the abnormal system behaviors.

The rest of the article is organized as follows. In Section 2, a preliminary overview of the operation of UDNs is presented. Section 3 first presents the PDE-based model for open-channel flow, followed by the reduced, data-driven system of the nonlinear ODE model obtained via spatial discretization. Then, the model of storage elements and the time-periodicity assumption on the disturbance signals are presented with the description of the system as a directed tree graph. In Section 4, the MHPE and MHSE techniques are detailed, whereupon Section 5 introduces the NMPC design and establishes the main control objectives. In Section 6, the results using data from a real-world network are presented. This is followed by Section 7 and Section 8, where a discussion, conclusions and future research directions are provided.

## 1.1 Nomenclature

Let $\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{m\times n}$, denote the field of real numbers, the set of real column vectors of length $n$ and the set of $m$ by $n$ real matrices, respectively. Throughout the paper, all quantities mentioned are real. We use boldface letters for sets, such as $\boldsymbol{s} = \{s_1, ..., s_n\}$, as well as for vectors $\boldsymbol{x} = [x_1, ..., x_n]^T \in \mathbb{R}^n$. The superscript $^\top$ denotes transposition, and the operators $<, \leq, =, >, \geq$ denote element-wise relations of vectors. Moreover, for a vector $\boldsymbol{x} \in \mathbb{R}^n$, $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$ denotes the Euclidean norm.

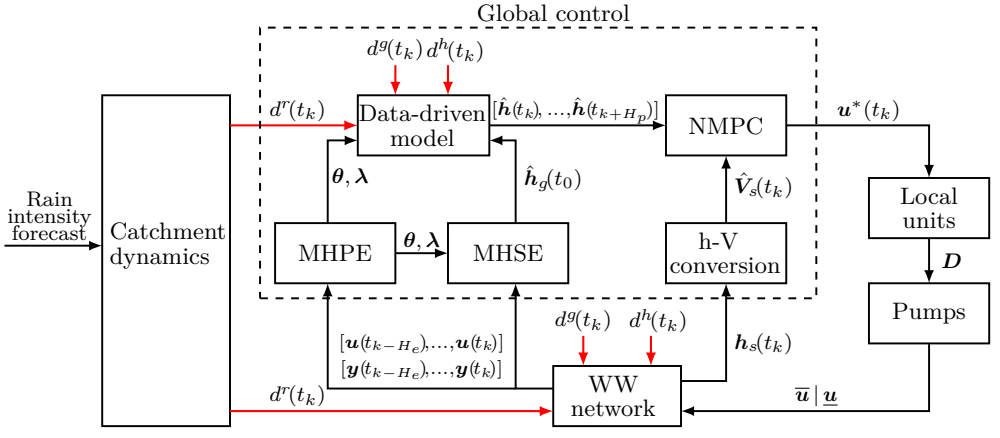


**Figure 1:** The proposed closed-loop control architecture, where rain intensities are known by means of weather forecasts and the transformation between the intensities and the runoff flow appearing in the sewers is characterized by the catchment dynamics. Red signals denoted with $d$ represent rain, household and groundwater inflow disturbances, $h$ and $V$ are the system states representing water level and water volume, respectively. Moreover, $u$ denotes the input of the aggregated flows which are delivered by locally-controlled pumps. The pipe network (plant) is represented by the WW (Waste Water) network block.

# 2 Drainage System overview

UDNs contain several elements, including gravitation pipes, manholes, pits, and in some cases, retention tanks. The most widely used actuators in pressurized sewer networks are pumps, typically installed inside wastewater pits where the sewage and rain are collected [Schütze et al., 2002]. These units often consist of one or several pumps in parallel, controlling the transport of the sewage from pit to pit. First, the water is pumped through a rising main, whereupon it gravitates through sewer pipes towards a downstream station, shown in Figure 2.

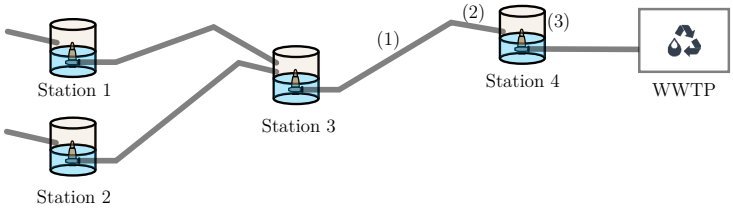


**Figure 2:** Tree topology of a pumped sewer network, where (1) illustrates rising mains, (2) gravity sewer pipes and (3) pumping stations [Butler and Davies, 2006].

UDNs typically have a tree structure, where the Waste Water Treatment Plant (WWTP) represents the root of the network.

# 3 System model

The modelling based on physics is introduced to show how the reduced model is obtained considering simple mass conservation rules and assumptions on the geometry of hydraulic structures. We aim to obtain a model structure with a low number of lumped parameters, where the system states are expressed by water levels. Besides, we show that the proposed internal model structure allows us to make assumptions on the initial parameters and their upper and lower bounds.

## 3.1 Physical transport model

Flow propagation in UDNs can be accurately computed by the full Saint-Venant (SV) equations, which are non-linear hyperbolic PDEs describing the mass and momentum of fluid:

$$\frac{\partial A_{x,t}}{\partial t} + \frac{\partial q_{x,t}}{\partial x} = \tilde{d}_{x,t}, \tag{1a}$$

$$\frac{\partial q_{x,t}}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q_{x,t}^2}{A_{x,t}}\right) + gA_{x,t}\Big(\overbrace{\frac{\partial h_{x,t}}{\partial x} + \underbrace{S_f - S_b}_{\text{Kinematic wave}}}^{\text{Diffusion wave}}\Big) = 0, \tag{1b}$$

where $q_{x,t}$ is the flow in the pipe and $\tilde{d}_{x,t} = d_{x,t}/dx$ represents lateral inflows per unit length, where $d_{x,t}$ is the lateral inflow hereinafter referred to as disturbance. $A_{x,t}$ is the wetted pipe area, $h_{x,t}$ represents the water level, furthermore $q_{x,t}, d_{x,t}, A_{x,t}$ and $h_{x,t}$ are functions from $(0, L) \times \mathbb{R}_+ \to \mathbb{R}_+$, where $L$ is the total length of the gravity pipe. The gravitational acceleration is denoted with $g$, moreover the slope term $S_b \in \mathbb{R}_+$ and friction term $S_f \in \mathbb{R}_+$ are assumed to be independent of $x$ and $t$, i.e. all pipe segments along the gravity pipe are modelled with assuming identical physical attributes.

The dynamics of each transport pipe in (1a) and (1b) are coupled through boundary conditions, hence the problem can become computationally demanding to solve in the case of complex networks [Xu et al., 2012]. Assumptions on the flow characteristics can lead to loss of dynamics, however, can lead to significant simplifications to the model structure. In this work, we utilize the Kinematic Wave approximation of the SV equations, thereby removing the left-hand-side terms of (1b). In this way, we omit the phenomena of wave attenuation, flow acceleration, and the phenomena of backwater effect[2]. These simplifications inherently mean that the considered flow characteristics are uniform and hence quasi-steady flow is assumed at all $x \in (0, L)$. The momentum equation in (1b) only considers two terms, i.e.,

$$S_b = S_f(q_{x,t}, h_{x,t}), \tag{2}$$

where the friction term $S_f$ is obtained from the Manning equation, which is an empirical formula for energy balance between gravity and friction, expressed by the level $h$ and flow $q$ variables [Schütze et al., 2002] as

$$S_b = \frac{n^2 q_{x,t}^2}{A_{x,t}^2 R_{x,t}^{4/3}}, \tag{3}$$

where $R = \frac{A}{P}$ is the hydraulic radius, $P \in \mathbb{R}_+$ is the wetted perimeter and $n \in \mathbb{R}_+$ is the Manning coefficient. Note that by knowing a map $f : A_{x,t} \mapsto h_{x,t}$, an expression between $q_{x,t}$ flow and $h_{x,t}$ level is constructed.

**Assumption 2.** *We assume a linear map $f$ between the wetted-area $A_{x,t}$ and water level $h_{x,t}$. It is assumed that semi-filled circular sewers are reasonably well-approximated by rectangular pipe shapes, i.e.,*

$$R_{x,t} = \frac{A_{x,t}}{P_{x,t}} \approx \frac{w h_{x,t}}{2h_{x,t} + w}, \tag{4}$$

*where $h_{x,t}$, $x \in (0, L)$ is the water level and the cross section is parametrized by the $w$ channel width shown in Figure 3.*

[2]Backwater occurs in sewers when the receiving water body becomes overloaded and therefore water volumes are accumulating at downstream of the connected hydraulic structure [Munier et al., 2008].

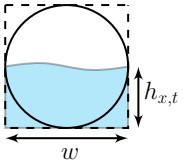


**Figure 3:** Semi-filled circular pipe approximated with rectangular geometry.

The independent variables remaining in the simplified SV PDEs in (1a) and (1b) are reduced to $q_{x,t}$ flow and $h_{x,t}$ level distributions on the domain $(0, L) \times \mathbb{R}_+$, given by:

$$w\frac{\partial h_{x,t}}{\partial t} + \frac{\partial q_{x,t}}{\partial x} = \frac{d_{x,t}}{\delta x}, \tag{5a}$$

$$q_{x,t} = \frac{\sqrt{S_b}}{n}\frac{(wh_{x,t})^{5/3}}{(2h_{x,t}+w)^{2/3}}, \tag{5b}$$

which is an approximation of the full dynamic SV-PDEs. Note that Assumption 1 on the pipe geometry means a linear scaling from levels $h_{x,t}$ to flows $q_{x,t}$, which leads to inaccuracy for circular pipe profiles. However, the assumption on the linear geometry profile keeps the model complexity low.

## 3.2 Reduced, data-driven transport model

In order to formulate the transport dynamics in a form more amenable to system identification, the spatial discretization of the approximated SV-PDEs in (5a) and (5b) is considered. The gravity pipes are partitioned into $N_x$ non-overlapping $\delta x$ segments of length, while the signals $h_{x,t}$, $q_{x,t}$ and $d_{x,t}$ are approximated as piece-wise constant functions of the spatial coordinate $x$, as shown in Figure 4.

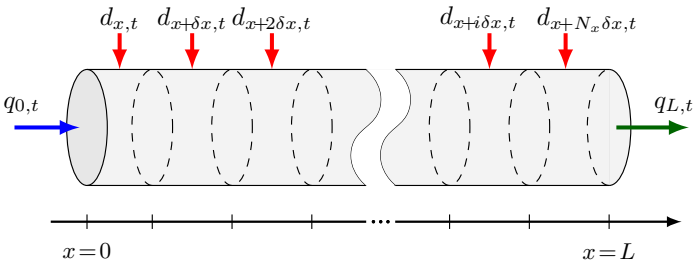


**Figure 4:** Gravity pipe divided into $N_x$, equal-sized, non-overlapping segments.

In Figure 4, $q_{0,t}$ and $q_{L,t}$ denote the flows corresponding to the upstream and downstream boundaries, respectively. Furthermore, $d_{x+i\delta x,t}$ represents the lateral inflows (disturbances) entering the $i^{th}$ pipe section, where $i \in \{1, 2, ..., N_x\}$.

**Remark 3.** *It is not necessary to partition gravity pipes into equal-sized $\delta x$ sections. The length of the spatial step $\delta x$ can be defined by the placement of manholes along the sewer pipes, among which some may be equipped with level sensors.*

**Remark 4.** *Close to the downstream end of gravity pipes ($x = L$), the discharge conditions of $q_{L,t}$ are influenced by the receiving hydraulic structure and the corresponding water levels [Roberson and Crowe, 1993], [Dey, 2002]. This relation imposes dynamics governing the water level $h_{L,t}$ in the last section. In this study, the effect of these types of dynamics are excluded, hence the positioning of water level sensors close to $x = L$ are chosen such that*

$$L - x \geq s_{min}, \tag{6}$$

*where $s_{min} \in \mathbb{R}_+$ denotes the minimal distance from the end of the channel where level sensors at position $x$ should be placed. The criteria of choosing $s_{min}$ for the free fall condition of fluids, based on the diameter of open-channel pipes is detailed in [Roberson and Crowe, 1993, pp.698-699].*

The spatial discretization of (5a) and (5b) is done by the backward Euler method. The left boundary (upstream) is defined at $x = 0$ and the right boundary (downstream) at $x = L$. Then, the SV-PDEs are reduced to the following system of finite dimensional ODEs:

$$\frac{dh_{x,t}}{dt} = \theta_1(q_{x-\delta x,t} - q_{x,t} + d_{x,t}), \quad \forall\ x \in (0, L), \tag{7a}$$

$$q_{x,t} = \theta_2 \frac{h_{x,t}^{5/3}}{(h_{x,t} + \theta_3)^{2/3}}, \quad \forall\ x \in (0, L), \tag{7b}$$

where the physical constants and the spatial time step are lumped into the parameters

$$\theta_1 \triangleq \frac{1}{w\delta x}, \quad \theta_2 \triangleq \frac{\sqrt{S_b} w^{\frac{5}{3}}}{2^{\frac{2}{3}} n}, \quad \theta_3 \triangleq \frac{w}{2}, \tag{8}$$

where $\theta_1, \theta_2, \theta_3 \in \mathbb{R}_+$. Note that $\theta_3$ is directly related to the width parameter $w$ and $\theta_1$ would change along the pipe in case of non-equal spatial steps $\delta x$. For the sake of simplicity, the model is presented with fixed $\delta x$ spatial steps.

**Remark 5.** *Due to the spatial discretization, numerical distortion is introduced in the traveling wave [Xu et al., 2012], which compensates for the flow attenuation phenomena in gravity pipes. This artificial attenuation vanishes as $\delta x \to 0$.*

In order to obtain the state equation with water levels as states, the section flow distribution $q_{x,t}$ in (7a) is substituted with water levels from (7b), which yields

$$\frac{dh_{x,t}}{dt} = \theta_1\theta_2\left(\frac{h_{x-\delta x,t}^{5/3}}{(h_{x-\delta x,t} + \theta_3)^{2/3}} - \frac{h_{x,t}^{5/3}}{(h_{x,t} + \theta_3)^{2/3}}\right) + \theta_1 d_{x,t}, \tag{9}$$

where, opposed to previous work in [Balla et al., 2020a], the flow balance in the SV equations is reformulated with physically measurable water levels. For ease of notation, let us define a non-linear map $g : \mathbb{R}_+ \to \mathbb{R}_+$ as

$$g : (h_{x,t}, \theta_3) \mapsto \frac{h_{x,t}^{5/3}}{(h_{x,t} + \theta_3)^{2/3}}, \quad \forall\, x \in (0, L). \tag{10}$$

Then, the transport flow model is completed by introducing the boundary conditions into the $N_x$ coupled ODEs, i.e.,

$$\begin{aligned}
\frac{dh_{0,t}}{dt} &= \theta_1(q_{0,t} + d_{0,t}) - \theta_1\theta_2 g(h_{\delta x,t}, \theta_3), \\
&\vdots \\
\frac{dh_{x,t}}{dt} &= \theta_1\theta_2\big(g(h_{x-\delta x,t}, \theta_3) - g(h_{x,t}, \theta_3)\big) + \theta_1 d_{x,t}, \\
&\vdots \\
\frac{dh_{L,t}}{dt} &= \theta_1\theta_2 g(h_{L-\delta x,t}, \theta_3) + \theta_1(d_{L,t} - q_{L,t}),
\end{aligned} \tag{11}$$

where $d_{x,t}$ is the unknown disturbances in form of lateral inflows. Besides, the upstream boundary flow $q_{0,t}$ is subject to control and hereinafter denoted as $u$. The downstream boundary flow $q_{L,t}$ is the discharged output, which we consider as the controlled model output, hereinafter denoted as

$$y = \theta_2 g(h_{L,t}, \theta_3), \tag{12}$$

where $h_{L,t}$ represents the water level at the downstream boundary $x = L$. Note that (12) is the parametric form of (3), relating the level to flow.

## 3.3 Disturbance model

The proposed model described in (11) aggregates the unknown, scaled disturbances in $d_{x,t}$. The disturbances are typically composed of several different types of inflows:

$$d_{x,t} \triangleq d^r_{x,t} + d^h_{x,t} + d^g_{x,t}, \quad \forall\, x \in (0, L), \tag{13}$$

where $d^r_{x,t}$ denotes rain runoff, $d^h_{x,t}$ is the household flow due to human activity and $d^g_{x,t}$ stands for groundwater.

**Assumption 3.** *The disturbance flow generated by households has an inherent periodicity, such that $d^h_{x,t} = d^h_{x,t+T}$, where $T$ typically corresponds to one day. Moreover, disturbances generated by groundwater infiltration fulfill the constraint $\sum_{i=1}^{N_x} d^g_{i,t} = N_x d^g_{j,t}$, $\forall\ j \in \{1, 2, ..., N_x\}$, i.e. uniformly distributed along the whole length of gravity sewer pipes.*

**Remark 6.** *Seasonality with different time periodicity have been considered (e.g. weeks, months) in [Livera et al., 2011], where methodologies such as Fourier models have been used to decompose seasonal components in a broad range of applications.*

Besides, the rain runoff $d^r_{x,t}$ is generated by the dynamics of catchments where the intensity of rain precipitation is typically provided by weather forecasts. Several works have been done on relating rain radar forecasts to actual runoff flow in UDNs, e.g., [Löwe et al., 2016], [Löwe et al., 2014], [Ma et al., 2018], [Chang et al., 2001].

For modelling the periodic household flows $d^h_{x,t}$ and the constant groundwater $d^g_{x,t}$ flows, Fourier series are utilized. For simplicity, let us assume that $d^r_{x,t} = 0,\ \forall\ x \in (0, L)$, i.e., assuming a dry-weather period. Then, the scaled disturbances in (11) are defined as:

$$\begin{aligned} \tilde{d}_{x,t} &\triangleq \tilde{d}^g_{x,t} + \tilde{d}^h_{x,t} \\ &\triangleq \lambda_0 + \sum_{j=1}^{k} \big(\lambda_{1j} \cos(j\omega t) + \lambda_{2j} \sin(j\omega t)\big), \end{aligned} \tag{14}$$

where the set of disturbance parameters is $\boldsymbol{\lambda} \triangleq \{\lambda_0, \lambda_{11}, \lambda_{21}, ..., \lambda_{1k}, \lambda_{2k}\} \in \mathbb{R}^{2k+1}$. The angular frequency $\omega$ corresponds to a period of one day and $k \geq 2$ is the number of frequency terms in the truncated Fourier series. The transport model in (11) and (12), in combination with the disturbance model in (14) are used to find parameters $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$.

## 3.4 Storage model

Stored volume within the network is represented through wastewater pits, among which some are specifically constructed to retent extreme peak flows caused by sudden rainfall runoffs. An example for such storage structure is shown in Figure 5. These pits are distinguished from single wastewater pits due to their large capacity and therefore referred to as retention pits in the rest of this paper. For each storage element at pumping station $i \in \{1, ..., N_s\}$, the infinitesimal level change is computed as the sum of all in- and outflows as

$$\frac{df_V(h_{s,i})}{dt} = d_{s,i} + \sum_{j=1}^{N_y} y_j - u_i, \tag{15}$$

where $d_s$ denotes disturbance inflows to storage tanks, $h_s$ is the water level in storage units and $u$ is the sum of controlled pump flows moving the water towards the next pumping station in line. Note that $u$ is equivalent to the inlet flow $q_{0,t}$ of a gravity pipe located between interconnected storage units, described in (11). Besides, $N_y$ is the number of gravity-driven transport links discharging to the $i^{th}$ storage unit and $y_j$ is the arriving discharge from the $j^{th}$ upstream pumping station, defined in (12). $N_s$ denotes the overall number of pumping stations in the UDN. Moreover, let us consider a map $f_V : \mathbb{R}_+ \to \mathbb{R}_+$ from water level $h$ to water volume $V$, where $f_V$ is strictly monotonic increasing.

**Assumption 4.** *For storage elements, $f_V(h)$ in (15) is approximated with a piece-wise linear, strictly monotonic increasing function, parameterized by the level-flow constant of storage tanks.*

Retention pits classify for the assumption on piece-wise linear behavior, while the relation between level and volume simplifies to linear in case of single pits. The piece-wise linear relation along with the hydraulic structure is shown in Figure 5.

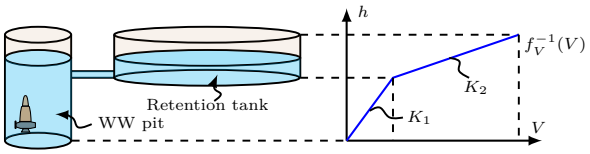


**Figure 5:** Level-Volume conversion for waste water pit with retention tank.

Tank constants $K_1$ and $K_2$ correspond to the slope of the $h$–$V$ conversion curve, where $K_2$ is only relevant if pits are equipped with retention tanks. Note that in dry weather the storage capacity of pits is sufficient, hence wastewater flow is typically bypassing the retention tank, acting as a single pit.

## 3.5 Network description

The links between system components define the topology of the network. The topology considered is a directed tree-graph with nodes representing storage (except the root) and edges transporting flow in between the nodes towards the root. The root of the graph is an outlet point, where the flow is discharged to the receiving environment, e.g., to the WWTP. The tree structure topology is shown in Figure 6.

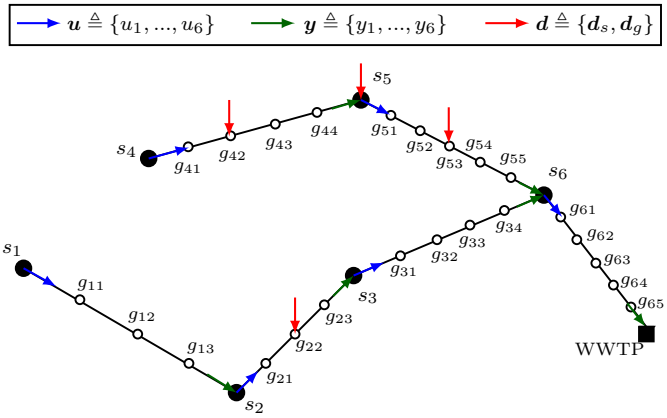


**Figure 6:** Graph representation where the filled nodes are pumping stations, empty nodes are manholes, whereas edges represent transport pipe segments.

Let us denote the set of nodes corresponding to tanks and pits at the pumping stations with $\mathcal{S} \triangleq \{s_i = (h_{s,i}, d_{s,i}, u_i) \,|\, i \in \{1, ..., N_s\}\}$, where $h_{s,i}$ is the water level, $d_{s,i}$ is the unknown flow disturbance defined in (13) and $u_i$ is the

controlled flow of the $i^{th}$ pumping station. The remaining nodes represent manholes along the gravity sewer transport links. These set of nodes are denoted with $\mathcal{G} \triangleq \{g_{ij} = (h_{g,ij}, d_{g,ij}) \,|\, i \in \{1, ..., N_s\}, j \in \{1, ..., N_g\}\}$, where $h_{g,ij}$ are the water levels in the $j^{th}$ segment of the gravity pipe rooting from the $i^{th}$ upstream station. Furthermore, $d_{g,ij}$ are lateral inflows along the $i^{th}$ gravity pipe, entering through the $j^{th}$ manhole. These disturbance components are given in (13). The numbering of manholes along the gravity pipes denotes the upstream pumping station (first digit) from which they are numbered in an increasing order towards the downstream station (second digit). The connections between the storage and junction nodes are defined by the piping layout.

Note that the set of $\mathcal{G}$ junction points also represents storage by means of the volume of pipe sections, as the spatially discretized and reduced SV-based model is equivalent to volumes connected in series, where the set of water levels $\{h_{g,i1}, ..., h_{g,iNg}\}, i \in \{1, ..., N_s\}$ relate to water volumes stored in each segment. However, we distinguish between $\mathcal{S}$ and $\mathcal{G}$ for the reason that nodes in $\mathcal{S}$ are subject to control $u$.

### 3.6 Discrete network model

In this study, discrete-time network dynamics are utilized for solving the MHE and NMPC problems. The transport and storage dynamics, described in Section 3.2 and Section 3.4, are given for each individual network element, respectively, as

$$\hat{\boldsymbol{h}}_g(t_{k+1}) = \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\lambda}}\big(u(t_k), \hat{\boldsymbol{h}}_g(t_k), \boldsymbol{d}_g(t_k)\big), \tag{16a}$$

$$\hat{h}_s(t_{k+1}) = \boldsymbol{H}\big(u(t_k), \hat{h}_s(t_k), d_s(t_k), \hat{\boldsymbol{y}}(t_k)\big), \tag{16b}$$

$$\hat{y}(t_k) = \boldsymbol{G}_{\boldsymbol{\theta}}\big(\hat{\boldsymbol{h}}_g(t_k)\big), \tag{16c}$$

where the numerical integration from $t_k$ to $t_{k+1}$ is done by the fixed step, $4^{th}$ order Runge-Kutta method. Moreover, $\hat{\boldsymbol{h}}_g(t_k) \in \mathbb{R}^{N_x}$ is the vector of water levels along a transport link between two pumping stations. The system dynamics corresponding to transport flows in (11) are defined by $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\lambda}} : \mathbb{R}_+ \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \to \mathbb{R}^{N_x}$. The discrete storage dynamics are given by $\boldsymbol{H} : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{N_y} \to \mathbb{R}_+$, where $N_y$ is the number of transport links discharging to the specific storage node. The outputs are represented by $\boldsymbol{G}_{\boldsymbol{\theta}} : \mathbb{R}^{N_x} \to \mathbb{R}$, corresponding to the discharged gravity flow previously described in (12).

## 4 Moving Horizon Estimation

In order to incorporate system knowledge in the state and parameter estimation in form of constraints, a MHE approach is utilized in this paper. Past data samples of the inputs, i.e., pump flows $\{u(t_{k-H_e}), u(t_{k-H_e+1}), ..., u(t_k)\}$ and the outputs, i.e., discharged gravity pipe flows $\{y(t_{k-H_e}), y(t_{k-H_e+1}), ..., y(t_k)\}$ are used up to the current time sample $t_k$, where $H_e$ is the estimation horizon.

Moreover, for each transport link $i \in \{1, ..., N_s\}$, let us define $\boldsymbol{h}_{g,i} \in \mathbb{R}^{N_x}$ as the vector of water levels in all $j \in \{1, ..., N_g\}$ pipe segments. The MHE problem regarding states (MHSE) and parameters (MHPE) is solved for each transport link $i$ individually. Therefore, we ease the notation and discard the $i$ and $j$ indices and we present the parameter and state estimation for a single transport link.

Additional outputs may be available by means of water level sensor measurements, placed in manholes along the main transporting sewers. Hence, we define $\boldsymbol{C} \in \mathbb{R}^{N_0 \times N_x}$ matrix, associated with a linear mapping which picks all the measured states. $N_0$ is the number of water level sensors along the transport link. Then, the output vector is given by

$$\boldsymbol{z}_g = \boldsymbol{C}\boldsymbol{h}_g + \nu, \tag{17}$$

where $\nu \in NID(0, \sigma^2)$ is white Gaussian noise accounting for measurement corruption, and $\boldsymbol{z}_g \in \mathbb{R}^{N_0}$. Past data samples $\{\boldsymbol{z}_g(t_{k-H_e}), \boldsymbol{z}_g(t_{k-H_e+1}), ..., \boldsymbol{z}_g(t_k)\}$ of these outputs are utilized together with the input $u$ and output $y$ flow data.

The main purpose of the MHPE is to identify the unknown dynamics of each transport link without using information about the physical properties of sewer pipes, such as pipe diameters, length, slope or roughness. Due to the linearized level-flow scaling introduced by Assumption 1 in Section 3.1, fixed model parameters might result in inaccurate flow predictions, based on whether the pipes are close to being filled or semi-filled. These characteristics can change over time due to seasonality, hence we utilize the MHPE method, attempting to adapt the model parameters to varying flow conditions. Moreover, the dynamics might change over time due to sludge accumulating within certain sections of the pipes, for which the proposed MHPE method is also able to account. As a natural extension, the states are also estimated in a moving horizon fashion (MHSE).

In the following, we distinguish between the horizons of parameter and state estimations. For parameter estimation, we denote the length of the horizon with $H_{pe}$ and for state estimation with $H_{se}$. Due to the slowly changing dynamics of sewer pipes, we argue that the MHPE is sufficient to carry out above the frequency of the NMPC, having at least one day up to a week long $H_{pe}$ horizon. However, the MHSE problem is executed with a minimum of one day long horizon and with the same frequency as the NMPC. The one day long MHSE horizon is due to the inherent periodicity of the waste water disturbance inflows $d^h$. Moreover, by calling the MHPE less frequent than the NMPC, we lower the typically high computation demand of MHE algorithms, where state and parameter estimations are carried out simultaneously [Allgöwer et al., 1999].

## 4.1 Parameter estimation

The MHPE problem of transport flows is formulated as a constrained, least-squares nonlinear minimization problem.

**Remark 7.** *The control inputs $u$, depicted in Fig. Figure 6, are estimated considering the polynomial expression of fixed-speed wastewater pumps in the form*

$$\hat{u} = s\mu_0 + s\mu_1 \Delta p + s\mu_2 P_p, \tag{18}$$

*where $s$ is the number of running pumps at the pumping station, $\Delta p$ is the differential pressure and $P_p$ is the power consumption of the pumps [Kallesøe and Knudsen, 2016]. Besides, the outputs $y$ corresponding to discharged flows in Fig. Figure 6 are estimated using mass conservation, detailed in [Kallesøe and Knudsen, 2016]. In this work, we use the outcome of the referenced flow estimation algorithm to provide outputs for the MHE problem.*

Let $\boldsymbol{\theta} \triangleq \{\theta_1, \theta_2, \theta_3\} \in \mathbb{R}_+$ denote the set of bounded system parameters and $\boldsymbol{\lambda} \in \mathbb{R}$ denote the parameters corresponding to the Fourier disturbance model. Then, for each transport link, the initial states $\hat{\boldsymbol{h}}_g(t_0)$, the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ are found by solving the following finite-dimensional constrained Nonlinear Programming (NLP) problem at time $t_k$:

$$\begin{aligned}\begin{pmatrix} \boldsymbol{\theta}^* \\ \boldsymbol{\lambda}^* \\ \hat{\boldsymbol{h}}_g^*(t_0) \end{pmatrix} = \underset{\boldsymbol{\theta},\boldsymbol{\lambda},\hat{\boldsymbol{h}}_g(t_0)}{\operatorname{argmin}} & \sum_{i=k-H_{pe}}^{k} \big(y(t_i) - \hat{y}(t_i)\big)^2 \\ & + W_1 \|\boldsymbol{z}_g(t_i) - \hat{\boldsymbol{z}}_g(t_i)\|^2,\end{aligned} \tag{19a}$$

subject to sewer dynamics:

$$\hat{\boldsymbol{h}}_g(t_{i+1}) = \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\lambda}}\big(u(t_i), \hat{\boldsymbol{h}}_g(t_i), \boldsymbol{d}_g(t_i)\big), \tag{19b}$$

$$\hat{y}(t_i) = \boldsymbol{G}_{\boldsymbol{\theta}}\big(\hat{\boldsymbol{h}}_g(t_i)\big), \tag{19c}$$

$$\hat{\boldsymbol{z}}_g(t_i) = \boldsymbol{C}\hat{\boldsymbol{h}}_g(t_i), \tag{19d}$$

and inequality constraints:

$$\mathbf{0} \leq \hat{\boldsymbol{h}}_g(t_i) \leq \overline{\boldsymbol{h}}_g, \tag{19e}$$

$$\mathbf{0} < \boldsymbol{\theta} \leq \overline{\boldsymbol{\theta}}, \tag{19f}$$

$$0 \leq \hat{y}(t_i) \leq \overline{y}, \tag{19g}$$

where $\hat{\boldsymbol{h}}_g(t_i) \in \mathbb{R}^{N_x}$ is the vector of states corresponding to a transport link. Note that $y$ represents the discharged flow, while $\boldsymbol{h}_g$ represents the vector of water levels in the manholes. Therefore, we use $W_1$ as a weighing constant in (19a), scaling the water levels to the magnitudes of the discharged flows. The constraints in (19e), (19f) and (19g) impose bounds on state variables, parameters and the output, respectively. Note that the states $\hat{\boldsymbol{h}}_g$ and the output variable $y$ correspond to physically measurable water levels and the discharged flow in sewer pipes, respectively. The water level measurements addressed in (17) are denoted by $\boldsymbol{z}_g$. The state and output bounds are chosen

consistent with physically meaningful values, such as water levels and flows are never negative inside the pipes.

Moreover, the upper bound on the states $\overline{\boldsymbol{h}}_g$ is the maximum allowed water level in manholes defined by the physical height from the bottom to the surface. From (8), we know that the pipe parameters $\boldsymbol{\theta}$ are positive. Besides, using (8), we approximate a physically meaningful maximum value for the pipe diameters, spatial step, time steps, and friction values. In (19g), the flow is assumed to be non-negative inside the pipe and the maximum is defined as the physically possible full-pipe flow. The bound constrained nonlinear minimization problem in (19) is then solved via a gradient descent algorithm.

Note that the number of sections $N_x$, illustrated in Figure 4, is treated as an auxiliary variable in the NLP, meaning that the MHPE problem can be carried out multiple times with different grid sizes to find the optimal number of sections regarding some performance index, e.g., Root Mean Squared Error (RMSE). This procedure is not detailed here, as the reader may consult a previous study focusing on how to choose grid size for a flow-based SV-PDE model in [Balla et al., 2020a].

## 4.2 State estimation

Full state measurement in the proposed sewer system model requires sensor installation inside all available manholes within the network. This is neither economically feasible, nor required by the control point of view. However, it is assumed that there is a subset of states $\boldsymbol{z}_g$ which are measured. Similarly to the MHPE problem in (19a), the full system states, i.e., $\boldsymbol{h}_g$ water levels are being reconstructed out of a few output measurements by means of the MHSE. However opposed to the MHPE problem in (19a), the state estimation is solved at each control time step $t_k$, thus providing initial state estimates for the NMPC. The MHSE reconstructs $\boldsymbol{h}_g(t_{k-H_{se}}), ..., \boldsymbol{h}_g(t_k)$ states, based on the measured inputs $\boldsymbol{u}(t_{k-H_{se}}), ..., \boldsymbol{u}(t_k)$, measured outputs $y(t_{k-H_{se}}), ..., y(t_k)$ and $\boldsymbol{z}_g(t_{k-H_{se}}), ..., \boldsymbol{z}_g(t_k)$ over the horizon $H_{se}$, while the dynamics are provided as constraints. The MHSE is defined by the following optimization problem:

$$\begin{pmatrix} \hat{\boldsymbol{h}}_g^*(t_{k-H_{se}}) \\ \vdots \\ \hat{\boldsymbol{h}}_g^*(t_k) \end{pmatrix} = \underset{\hat{\boldsymbol{h}}_g(t_{k-H_{se}}),\ldots,\hat{\boldsymbol{h}}_g(t_k)}{\operatorname{argmin}} \sum_{i=k-H_{se}}^{k} \left(y(t_i) - \hat{y}(t_i)\right)^2 + W_2 \|\boldsymbol{z}_g(t_i) - \hat{\boldsymbol{z}}_g(t_i)\|^2, \tag{20}$$

subject to the dynamics in (19b) to (19d), the state constraint in (19e) and the output constraint in (19g). $W_2$ is a weighing matrix for scaling levels to flows, similarly as in (19a). Note that from the solution of the MHSE problem in (20), the estimated state vector at the current time step $\hat{\boldsymbol{h}}_g^*(t_k)$ is used. The same gradient descent algorithm [Wills and Ninness, 2008] is used to solve the problem in (20), as for the MHPE.

# 5 Control design

The vector of control variables is defined by $\boldsymbol{u} \in \mathbb{R}^{N_s}$, where all individual pump flows are aggregated at the nodes $s \in \mathcal{S}$, representing the $N_s$ pumping stations in the network. The states correspond to levels along transport links and levels in storage units, e.g., pits. The state vector is defined by:

$$\boldsymbol{h} \triangleq (\boldsymbol{h}_s^\top, \boldsymbol{h}_{g,1}^\top, \boldsymbol{h}_{g,2}^\top, ..., \boldsymbol{h}_{g,N_s}^\top)^\top, \tag{21}$$

where $\boldsymbol{h}_s \in \mathbb{R}^{N_s}$ represents the vector of levels in storage elements and for each $i \in \{1, ..., N_s\}$ transport link $\boldsymbol{h}_{g,i} \in \mathbb{R}^{N_{x,i}}$ consists of $N_{x,i}$ entries depending on how many sections each transport link is discretized into. The outputs, i.e., discharged flows at the end of each transport link, are given by:

$$\boldsymbol{y} \triangleq (y_1, y_2, ..., y_{N_s})^\top, \tag{22}$$

where the last element $y_{N_s}$ is the discharged flow leading to the root of the network, which we hereinafter denote as $y_w$. Introduced previously, the closed-loop control scheme together with the MHE problem is depicted in Figure 1. Note that the rain run-off dynamics along with the weather forecasts provide flow inputs to the proposed closed-loop control scheme.

## 5.1 NMPC problem

To account for both the dry- and wet-weather loads in a computationally efficient way, the NMPC problem is formulated over two subsequent prediction horizons. To this end, let $H_{p_1}$ denote the predictions over the near future (nowcasts) and $H_{p_2}$ the predictions further in the future (forecasts), respectively. This formulation of the NMPC problem is motivated by the inherent periodicity of the household disturbances $d^h$, which typically corresponds to one day. However, the network is exposed to large disturbance loads in terms of the $d_r$ rain run-off, where the so-called nowcasts are reliable only within a short horizon. According to [Löwe et al., 2014], rainfall radars can provide sufficient accuracy of spatial and temporal resolution for urban catchments only up to a 2 (h) horizon. Therefore, computing the decision variables for $T = 24$ (h) is unnecessary, and results in high computational costs. Instead, let $H_p = H_{p_1} + H_{p_2}$ be the entire length of the horizon, $T_s$ the time step and let us define $\boldsymbol{h} \triangleq [\boldsymbol{h}_s^\top, \boldsymbol{h}_g^\top]^\top$ as the entire state vector. Then, the NMPC problem for the entire network is given as

$$\min_{\substack{\Delta \boldsymbol{u}(0),\dots,\Delta \boldsymbol{u}(H_{p_1}) \\ \Delta \boldsymbol{u}(H_{p_1}+1),\dots,\Delta \boldsymbol{u}(H_p-1)}} \sum_{k=0}^{H_p-1} \mathcal{L}\big(\Delta \boldsymbol{u}(t_k), \boldsymbol{h}(t_k), y_w(t_k)\big) + \mathcal{S}\big(\boldsymbol{h}(t_{H_p})\big) \tag{23a}$$

subject to transport link dynamics

$$\boldsymbol{h}_g(t_{k+1}) = \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\lambda}}\big(\boldsymbol{u}(t_k), \boldsymbol{h}_g(t_k), \boldsymbol{d}_g(t_k)\big), \tag{23b}$$

$$\boldsymbol{u}(t_{k+1}) = \boldsymbol{u}(t_k) + \Delta \boldsymbol{u}(t_k), \tag{23c}$$

$$y_w(t_k) = \boldsymbol{G}_{\boldsymbol{\theta}}\big(\boldsymbol{h}_g(t_k)\big), \tag{23d}$$

storage dynamics

$$\boldsymbol{h}_s(t_{k+1}) = \boldsymbol{H}\big(\boldsymbol{u}(t_k), \boldsymbol{h}_s(t_k), \boldsymbol{d}_s(t_k), \boldsymbol{y}(t_k)\big), \tag{23e}$$

state, input and output constraints

$$\underline{\boldsymbol{V}} + \boldsymbol{V}_{of}(t_k) \le \boldsymbol{f}_V\big(\boldsymbol{h}_s(t_k)\big) \le \overline{\boldsymbol{V}} + \boldsymbol{V}_{of}(t_k), \tag{23f}$$

$$0 \le \boldsymbol{h}_g(t_k) \le \overline{\boldsymbol{h}}_g, \tag{23g}$$

$$\underline{\boldsymbol{u}} \le \boldsymbol{u}(t_k) \le \overline{\boldsymbol{u}}, \tag{23h}$$

$$0 \le y_w(t_k) \le \overline{y}_w, \tag{23i}$$

terminal constraint

$$\underline{\boldsymbol{V}} \le \boldsymbol{f}_V\big(\boldsymbol{h}_s(t_{H_p})\big) \le \overline{\boldsymbol{V}}, \tag{23j}$$

$$0 \le \boldsymbol{h}_g(t_{H_p}) \le \overline{\boldsymbol{h}}_g, \tag{23k}$$

where $\Delta \boldsymbol{u}(t_k) \triangleq \boldsymbol{u}(t_{k+1}) - \boldsymbol{u}(t_k)$ is the input change. The integral action accounts for smooth and slow system response, avoiding sudden jumps in the control action. The proposed optimization problem in (23a) is solved for $[\Delta \boldsymbol{u}^\top(0), ..., \Delta \boldsymbol{u}^\top(H_{p_1})]^\top \in \mathbb{R}^{H_{p_1}}$, whereas the problem of finding the decision variables over $H_{p_2}$ is reduced to finding $H_{p2}/\tau$ number of optimization variables, where $\tau$ defines how many $T_s$ control steps each decision variable is kept constant. This is due to the fact that the control over $H_{p_2}$ does not require the same precision as for the nowcasts over $H_{p_1}$. The stage and terminal costs formulated in (23a) are sums of square-type functions, and the multiple operational objectives in the stage cost $\mathcal{L}$ are detailed in Section 5.2.

The dynamics $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\lambda}}$, $\boldsymbol{H}$ and $\boldsymbol{G}_{\boldsymbol{\theta}}$ are defined in (16a), (16b), (16c) for the entire network and the output equation in (23d) is formulated on the discharged flow $y_w$ arriving to the root of the network. The nonlinear level to volume conversion is kept outside of the optimization, where $\boldsymbol{f}_V$ is a piece-wise linear map from (15). Furthermore, the control is subject to state constraints on pipe states in (23g) and storage states in (23f), where $\boldsymbol{V}_{of} \in \mathbb{R}^{N_s}$ is the vector of slack variables, lifting the upper and lower state bounds. This variable is

considered as a virtual volume triggered at times when the physical limits of storage elements are extended. In case of an overflow, the slack variable lifts both the lower $\underline{\boldsymbol{V}}$ and upper $\overline{\boldsymbol{V}}$ state bounds, thereby keeping track of the excess storage [Gelormino and Ricker, 1994]. The upper bound of states corresponds to the physically maximum volume capacity in the storage nodes. The lower limit is defined by the user with the criteria that a minimum volume of water needs to be kept in the storage tanks at all times to fully cover the wastewater pumps, hence avoiding the dry-run of pumps.

**Remark 8.** *In case of overflows, the excess water volume leaves the network immediately. This is assured by constraining the slack variables*

$$\mathbf{0} \leq \boldsymbol{V}_{of}(t_k), \tag{24}$$

*meaning that spilled sewage escapes all $s \in \mathcal{S}$ storage nodes.*

Furthermore, (23h) imposes physical bounds on the minimum and maximum flow capacity of pumps. Equation (23i) formulates a constraint regarding the inflow capacity of the WWTP, where $\overline{y}_w$ is the maximum allowed inflow defined by the size of the WWTP. For closed-loop stability considerations of the NMPC, the terminal constrains in (23j) and (23k) are introduced along with the terminal cost $\mathcal{S}$ in (23a) to enforce stability [Allgöwer et al., 1999]. The formulation in (23a) is solved via a gradient descent algorithm, where the dynamics are discretized according to Section 3.6.

## 5.2 Objectives

The control problem addressed in (23) has multiple objectives with different priorities. For an extensive analysis on choosing objectives in UDN control, consult [Ocampo-Martinez, 2010; Mollerup et al., 2016]. To prioritize objectives, the stage cost is formulated as a linearly weighted sum and the terminal cost is given as

$$\mathcal{L}\big(\Delta \boldsymbol{u}(t_k), \boldsymbol{h}(t_k), y_w(t_k)\big) \triangleq \sum_{j=1}^{\Gamma} \lambda_j \mu_j \mathcal{F}_j(t_k), \tag{25a}$$

$$\mathcal{S}\big(\boldsymbol{h}(t_{H_p})\big) \triangleq \boldsymbol{h}^{\top}(t_{H_p}) \boldsymbol{P} \boldsymbol{h}(t_{H_p}), \tag{25b}$$

where $\lambda_j$ denotes the scaling weights among the different objectives and $\Gamma$ is the total number of the control objectives. The scaling constants $\mu_j$ normalize each objective term to dimensionless values, such that water levels and flows become comparable. Furthermore, the terminal cost $\mathcal{S}$ is defined for all states, where the symmetric positive definite matrix $\boldsymbol{P}$ is the solution to the associated Ricatti equation. Note that $\boldsymbol{P}$ is designed based on the weights $\lambda_j$ on the state and input terms in the stage cost function $\mathcal{L}$. Moreover, the Jacobian linearization of the network model is considered at an operating point, where state values are at their 25 % utilization of their upper limit. Furthermore,

disturbance and input flows are considered at the daily mean of household wastewater production without rain.

The most common control criteria in sewer network control is related to overflow minimization

$$\mathcal{F}_1(t_k) \triangleq \boldsymbol{V}_{of}^{\top}(t_k)\boldsymbol{\Omega}_1\boldsymbol{V}_{of}(t_k), \tag{26}$$

where $\boldsymbol{V}_{of} \in \mathbb{R}^{Ns}$ is the vector of slack variables, representing overflow volumes. The overflows $\boldsymbol{V}_{of}$ between stations are prioritized according to the diagonal $\boldsymbol{\Omega}_1$ matrix, where $\text{diag}\,(\boldsymbol{\Omega}_1) \in [0, 1]$. Note that the weight corresponding to the overflow objective $\lambda_1$ is significantly higher than any other weights, in order to make the use of the overflow slack variables undesirable if possible.

The penalty on water level in storage elements is given by

$$\mathcal{F}_2(t_k) \triangleq \boldsymbol{h}_s^{\top}(t_k)\boldsymbol{\Omega}_2\boldsymbol{h}_s(t_k), \tag{27}$$

where $\boldsymbol{h}_s \in \mathbb{R}^{Ns}$ is the vector of water levels in storage nodes and $\boldsymbol{\Omega}_2$ is the diagonal weighting matrix, where $\text{diag}\,(\boldsymbol{\Omega}_2) \in [0, 1]$. The level in storage nodes is minimized to avoid long retention times and thus odor problems occurring in the waste water tanks. Moreover, the weight matrix $\boldsymbol{\Omega}_2$ allows to adjust the filling sensitivity of storage elements, meaning that sensitive tanks are filled slower and emptied faster than less sensitive storage tanks.

The inputs are minimized such that

$$\mathcal{F}_3(t_k) \triangleq \Delta\boldsymbol{u}^{\top}(t_k)\boldsymbol{\Omega}_3\Delta\boldsymbol{u}(t_k), \tag{28}$$

where $\Delta\boldsymbol{u} \in \mathbb{R}^{Ns}$ is the vector of input change regarding the aggregated flows delivered by sewer pumps placed at each network node $s \in \mathcal{S}$. Moreover, $\boldsymbol{\Omega}_3$ is the weighting matrix between the network nodes, where $\text{diag}\,(\boldsymbol{\Omega}_3) \in [0, 1]$.

The system states in any $g \in \mathcal{G}$ nodes, i.e., gravity pipe sections are water levels, representing storage along the edges of the underlying network graph. Hence, we penalize manholes prone to suffer overflows:

$$\mathcal{F}_4(t_k) \triangleq \boldsymbol{h}_g^{\top}(t_k)\boldsymbol{\Omega}_4\boldsymbol{h}_g(t_k), \tag{29}$$

where $\boldsymbol{h}_g$ is the vector consisting of selected network nodes which can overflow under high loads. Similarly to all objectives, $\text{diag}\,(\boldsymbol{\Omega}_4) \in [0, 1]$ allows to adjust priority of overflows and filling sensitivity of manholes.

In this work, we consider the objective of controlling the inflow to the WWTP, which is formulated as follows

$$\mathcal{F}_5(t_k) \triangleq \Big(y_w(t_k) - \frac{1}{H_p}\sum_{k=0}^{H_p-1} y_w(t_k)\Big)^2, \tag{30}$$

where the inflow variation to the WWTP is minimized. This is achieved, by calculating a reference flow as an average inflow over the same time period as the time periodicity of the $d^h$ houshold disturbances, which typically corresponds

to one day. This objective allows to correct the irregular inflow pattern to the root of the network, which influences negatively the operation of the WWTP. An extensive study on the regulation of inlet flow to the WWTP is detailed in [van Heeringen et al., 2016].

# 6 Numerical results

We now present the numerical results. The results are related to the closed-loop control scheme performance when both the MHE and NMPC are considered. As presented in Section 3.6, the network model uses the fixed step, $4^{th}$ order Runge-Kutta method for the finite-difference approximation of the derivative terms. The optimization problem related to the MHSE and MHPE has been solved via a Gauss-Newton gradient-based method. This solver is chosen due to the reliable estimate of the Hessian for least-squares type problems, such as the MHE formulation in this paper [Wills and Ninness, 2008]. Furthermore, the optimization problem related to the NMPC has been solved via direct multiple shooting in the symbolic framework `CasADI` [Andersson et al., 2019]. A primer-dual interior point solver `IPOPT` [Wächter and Biegler, 2006] has been chosen to solve the nonlinear optimization problem in (23), due to its ability to leverage sparse linear algebra computations. Since the sampling interval is significantly short compared to the dynamics and the sampling time of the NMPC, the optimization problem has been solved by warm-starting at each control time step. Error tolerance of $10^{-5}$ has been chosen in both the MHE and NMPC problems. Moreover, all the numerical experiments have been carried out on a 2.6 (GHz), Intel Core i7 machine with 16-GB RAM.

Following the model methodology discussed in Section 3, the control-oriented model is identified based on measurements extracted from a physically-based HiFi network, shown in Figure 7 (left). The topological representation as a directed graph along with the location of sensors are depicted in Figure 7 (right). To test the NMPC with the MHE strategy, real rain intensity and wastewater flow are utilized starting from 1 September 2019 to 30 September 2019[3]. These data are used as the load to the HiFi case study network.

## 6.1 Baseline controller

In this work, we follow the guidelines proposed in [Lund et al., 2018] to benchmark the MPC performance, where the current state-of-art uses CSO and flooded volume as an evaluation measure. The proposed NMPC/MHE strategy is tested against an on/off rule-based controller, most commonly used as baseline control in both practice and literature [Lund et al., 2018], [García et al., 2015].

[3]Rain intensity data have been extracted from the weather archive of the Danish Meteorological Institute (DMI), while the domestic wastewater flow measurement data have been obtained and scaled down from the municipality of Fredericia, Denmark.

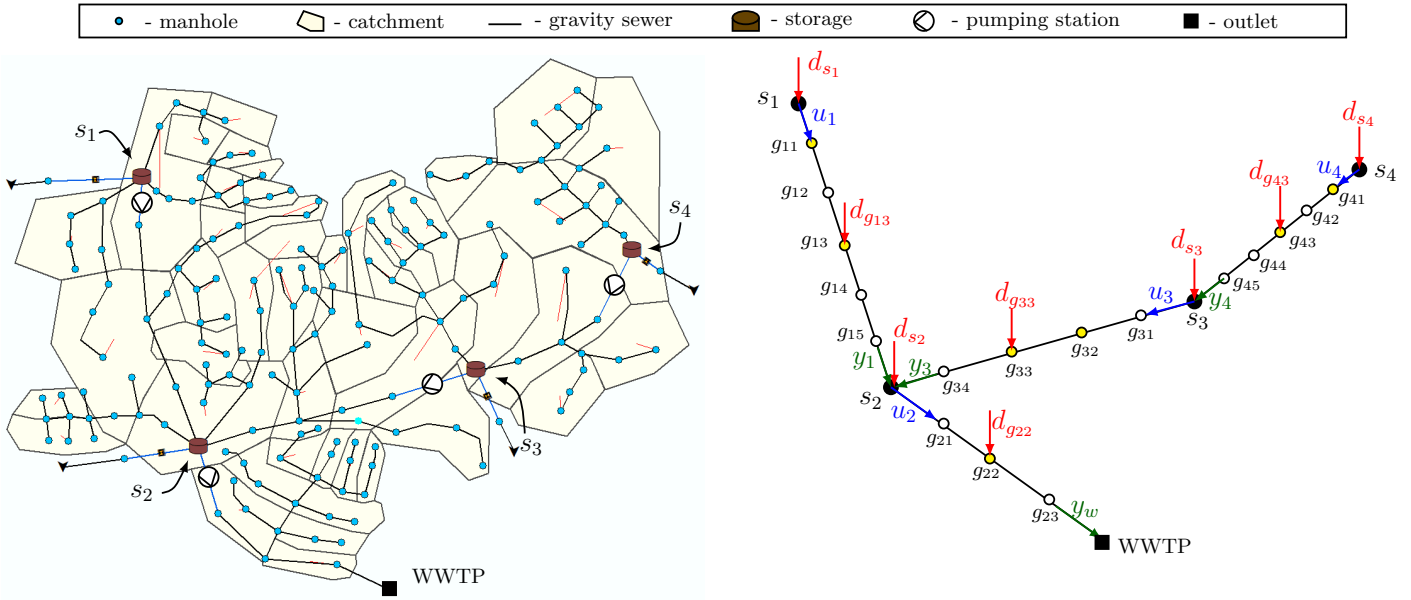


**Figure 7:** High-fidelity network model in the MIKE Urban simulation software (left), and the graph representation (right), where the number of empty nodes represents the number of discretized sections. The filled, yellow nodes represent level sensors placed in manholes.

The switching rule together with the aggregated flow provided by the pumps at each pumping station under the rule-based control is given by

$$u(t_k) = \begin{cases} \overline{u}, & \text{if} \quad h_s(t_k) \geq \overline{h}_s, \quad \forall t_k, \\ \underline{u}, & \text{if} \quad h_s(t_k) \leq \underline{h}_s, \quad \forall t_k, \\ u(t_{k-1}), & \text{otherwise}, \qquad \forall t_k, \end{cases} \tag{31}$$

where $h_s(t_k)$ is the measured water level in the storage element. Upper and lower bounds of the inlet flow $\overline{u}, \underline{u}$ are equivalent to the bounds in (23h), corresponding to the maximum and minimum flow capacity of the pumps. Threshold values $\overline{h}_s$, $\underline{h}_s$ are equivalent to (23f).

## 6.2 Case study

The topological properties of the HiFi network shown in Figure 7 are summarized in Table 1. We consider a combined sewer network, where both rain runoff

**Table 1:** MIKE Urban HiFi simulation properties.

| Attribute | Number | Variable | Unit |
|---|---|---|---|
| Single pits | 3 | $h_s$ | (m) |
| Retention pits | 1 | $h_s$ | (m) |
| Pumping stations | 4 | $u$ | ($m^3$/h) or ($m^3$/s) |
| Level sensors in manholes | 7 | $h_g$ | (m) |
| Catchment runoff | 45 | $d^r$ | ($m^3$/h) or ($m^3$/s) |
| Waste water inflow | 10 | $d^h$ | ($m^3$/h) or ($m^3$/s) |
| Treatment plants | 1 | $y_w$ | ($m^3$/h) or ($m^3$/s) |

and wastewater enters the sewer via the catchments (yellow areas) and the manholes (junction points), respectively. The network consists of 170 manholes, 170 gravity pipes, moreover three single pits ($s_1, s_3$ and $s_4$) and a retention pit ($s_2$). Using the proposed modelling methodology, the tree graph representation of the UDN and the control variables in the reduced graph representation are given by:

$$\boldsymbol{u} \triangleq (u_1, u_2, u_3, u_4)^\top, \tag{32a}$$

$$\boldsymbol{h} \triangleq (h_{s_1}, h_{s_2}, h_{s_3}, h_{s_4}, \boldsymbol{h}_{g_1}^\top, \boldsymbol{h}_{g_2}^\top, \boldsymbol{h}_{g_3}^\top, \boldsymbol{h}_{g_4}^\top)^\top, \tag{32b}$$

$$\boldsymbol{y} \triangleq (y_1, y_2, y_3, y_w)^\top, \tag{32c}$$

$$\boldsymbol{d} \triangleq (d_{s_1}, d_{s_2}, d_{s_3}, d_{s_4}, d_{g_{13}}, d_{g_{22}}, d_{g_{33}}, d_{g_{43}})^\top, \tag{32d}$$

where the state vector $\boldsymbol{h}$ consists of the gravity pipe subvector states $\boldsymbol{h}_{g_1} \in \mathbb{R}^5$, $\boldsymbol{h}_{g_2} \in \mathbb{R}^3$, $\boldsymbol{h}_{g_3} \in \mathbb{R}^4$, $\boldsymbol{h}_{g_4} \in \mathbb{R}^5$. Moreover, the number of pumping stations is $N_s = 4$ and the rain and domestic wastewater disturbances are concentrated on certain network nodes. The control time step of the NMPC is $T_{\text{NMPC}} =$

10 (min), while the rule-based controllers operate with a $T_{\text{On/Off}} = 1$ (min) sampling period. The prediction horizon for nowcasts is $H_{p_1} = 2$ (hours) and for the forecasts $H_{p_2} = 22$ (hours), summing up to a total of one-day horizon.

The MHPE is carried out with a horizon $H_{pe} = 2$ (day) and utilized with a $T_{pe} = 6$ (hours) period time. At every 6 (hours), the MHPE uses data from the past two days and updates the $\boldsymbol{\theta}$ system and $\boldsymbol{\lambda}$ disturbance parameters accordingly. A minimum of two days has been chosen to detect the one-day periodicity of the household wastewater with the Fourier disturbance model. The MHSE is carried out with the same horizon as the MHPE, i.e., $H_{se} = 2$ (day) and utilized with the same frequency as the NMPC, i.e., $T_{se} = 10$ (min).

## 6.3 Simulation environment

To test the NMPC/MHE controller, the MIKE Urban [MIKE powered by DHI, 2017] simulation software has been used to simulate the HiFi network model depicted in Figure 7. MIKE Urban allows for the hydraulic and hydrodynamic simulation of flows and water levels by numerically solving the full SV equations in (1). The model of the network in MIKE Urban is defined by the true physical parameters of the hydraulic components. In Figure 7, the catchments (yellow areas) are connected to manholes, hence water volumes enter the pipe network through the network nodes. The simulation is done in two steps: First, the network loads are computed with the catchment dynamics. Then, the rain runoff together with the household waste and groundwater appears as a load (marked with red arrows in Figure 7).

In this work, the NMPC/MHE strategy is used as an upper level controller, where the MIKE Urban model is simulated as a virtual reality. To this end, we utilize the MIKE 1D Application Programming Interface (API) [MIKE powered by DHI, 2019], [MIKE powered by DHI] which allows us for setting flow references to the pumps and reading flow and level values of hydraulic structures during simulation. These flow references are calculated at every $T_{\text{NMPC}}$ time and then used as set-points for local PID controllers based on (virtual) flow sensor measurements placed right after the downstream end of the pumps. The HiFi model runs with a sampling time of $T_{\text{on/off}}$, however the set-points for the PIDs are kept constant during the time interval $T_{\text{NMPC}}$.

## 6.4 Identification results

To estimate the parameters and the initial states in transport pipes, the measurements $\boldsymbol{z}$ along with the historical data on the estimated inlet and discharged flows $\boldsymbol{u}$ and $\boldsymbol{y}$ are utilized. To show the capabilities of the MHE approach, the initial conditions estimated for the problem in (23) are compared to the measurements in the HiFi simulation, shown in Fig Figure 8.

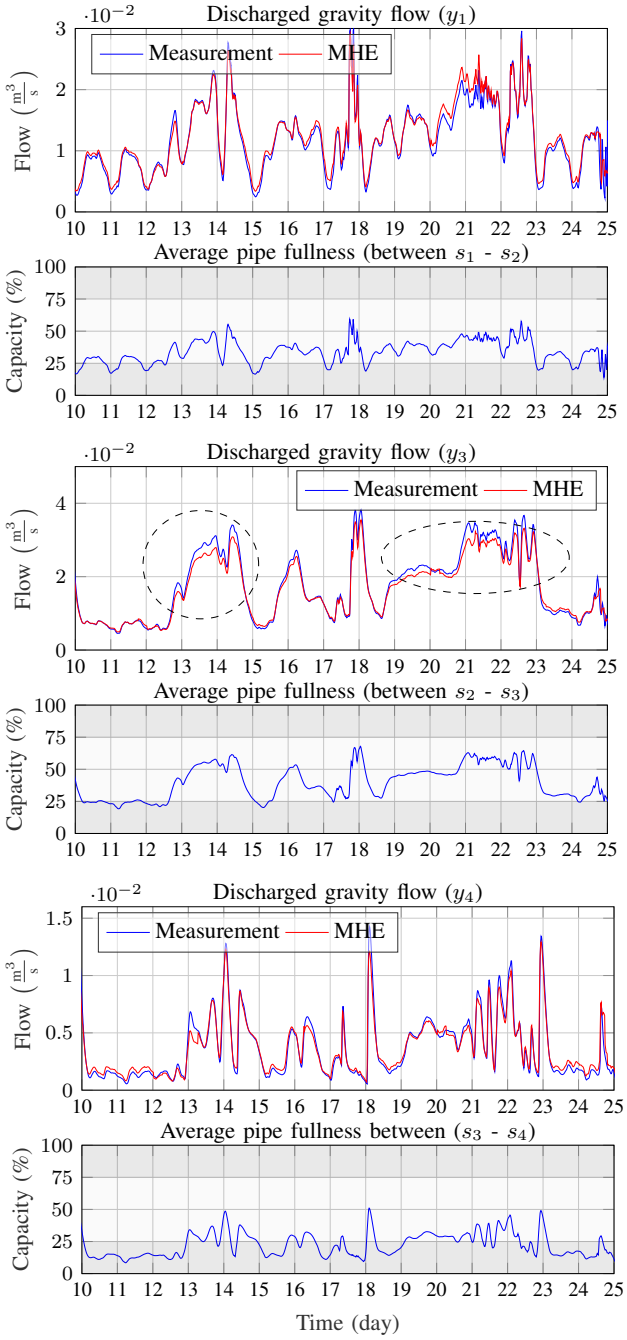


**Figure 8:** One-step prediction of $y_1$, $y_2$ and $y_3$ gravity flows.

The results show 15 days, where the estimated flow $\hat{y}(t_0)$ is the result of the MHPE and MHSE blocks depicted in Figure 1. Note that instead of showing the estimated states $\hat{h}_g(t_0)$, we rather show the discharged flow $\hat{y}(t_0)$, obtained by (19c). From the application point of view, this is reasonable, since the water level in the last section of a transport pipe does not indicate how the volume is affected in the receiving hydraulic structure (storage tanks), opposed to the volumetric flow rate. Besides, for each transport link $g_1, g_2$ and $g_3$, we show the average pipe fullness along $x \in L$ length of the pipes, indicating the capacity of each pipeline.

In the HiFi simulation environment, all sewer pipes are circular, hence the flow-level translation imposed by Assumption 1 (Section 3) is only accurate for small variations of water level. In order to show the variations of water levels inside gravity pipes, we illustrate two different operating regions (shaded areas in Fig. 8.). The middle range of the pipe is defined between 25 – 75 %, moreover the lower and upper regions between 0 – 25 % and 75 – 100 %, respectively. Small level variations within these regions are expected to yield accurate flow estimates based on Assumption 1.

As shown in Fig. Figure 8, the one-step predictions of the MHE strategy produces accurate estimates of the discharged flows $\hat{y}(t_0)$ in comparison with the flow measurements ($y(t_0)$) from the HiFi model. This is achieved without using any flow sensor in the network, however, assuming the linear flow-level relation in the internal model. The prediction results in Figure 8 show inaccurate flow estimates at certain time steps imposed by the simplified pipe geometries. This is because the internal model with the simplified geometry attempts to produce flows close to the ones obtained by the linear flow-level mapping, rather than the actual flow. This is most visible on $y_3$ at periods encircled with dashed black lines. During both of these periods, the pipes are filled up from 25 % to 50 %, where the previous level with the time window of the MHPE only provides information of low-filled, slow-varying level conditions. Hence, the internal model underestimates the actual flow by calculating lower volume than there is inside the middle operating range of circular pipes.

## 6.5 Control results

The control results aim to show the benefits of distributing level sensors along the network to obtain the data-driven network model from the full SV-PDEs. The proposed methodology is compared with a traditional, two-point controller detailed in Section 6.1, most commonly used by water utility operators. The NMPC acts as a global controller and computes reference points to local controllers (as depicted in Figure 1). To evaluate the closed-loop performance of the NMPC/MHE strategy, we selected two days with heavy overflows due to the insufficient capacity in the network. The numerical results are shown in Figure 9-A and Figure 9-B for each $i \in \{1, ..., N_s\}$ pumping station, showing the time evaluation of the disturbances, overflows, tank levels and the pumped inlet flows.

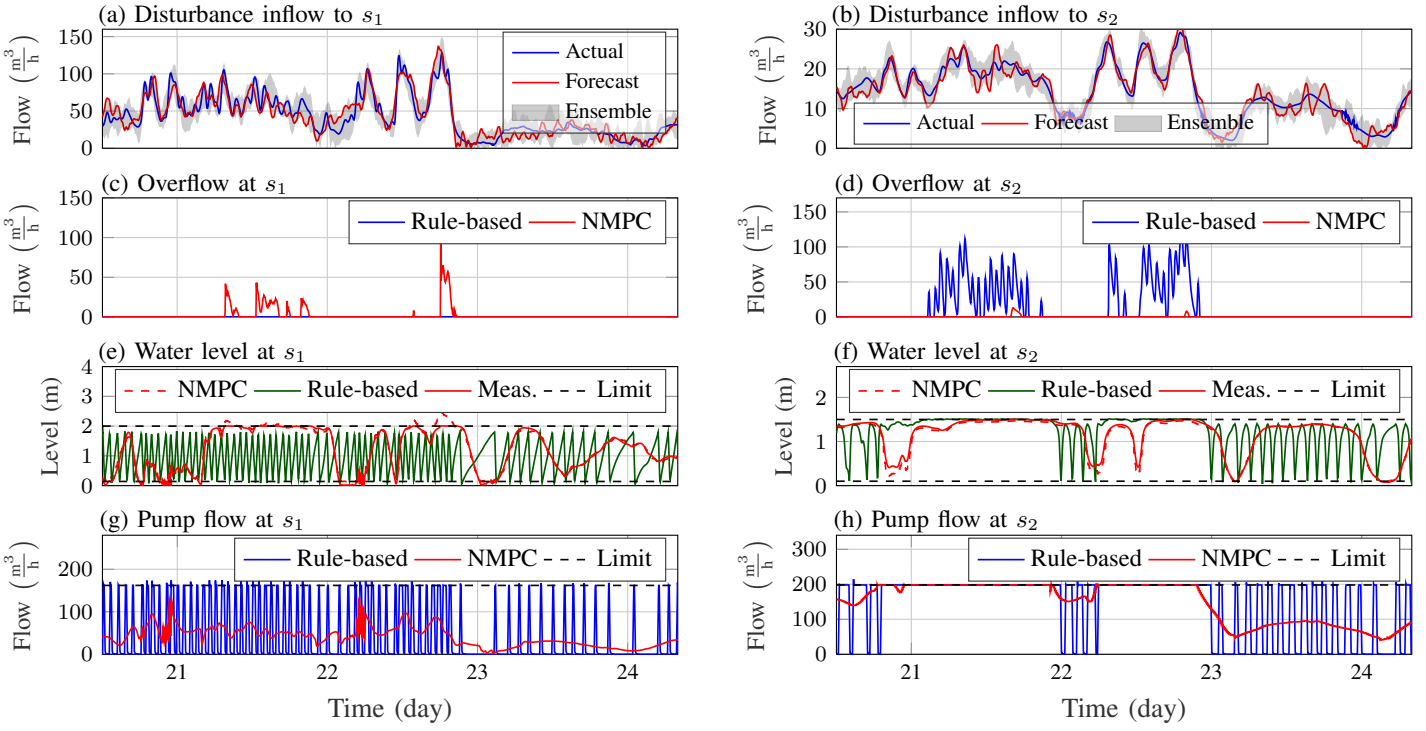


**Figure 9-A:** NMPC with MHPE and MHSE results compared with rule-based control. (a), (b), (i), (j) show disturbance inflows $d_{s,i}$, (c), (d), (k), (l) show overflows $q^{of}_{s,i}$, (e), (f), (m), (n) show states (water levels), while (f), (h), (o), (p) show inputs (pump flows) for each $i \in \{1, ..., N_s\}$ station, respectively.

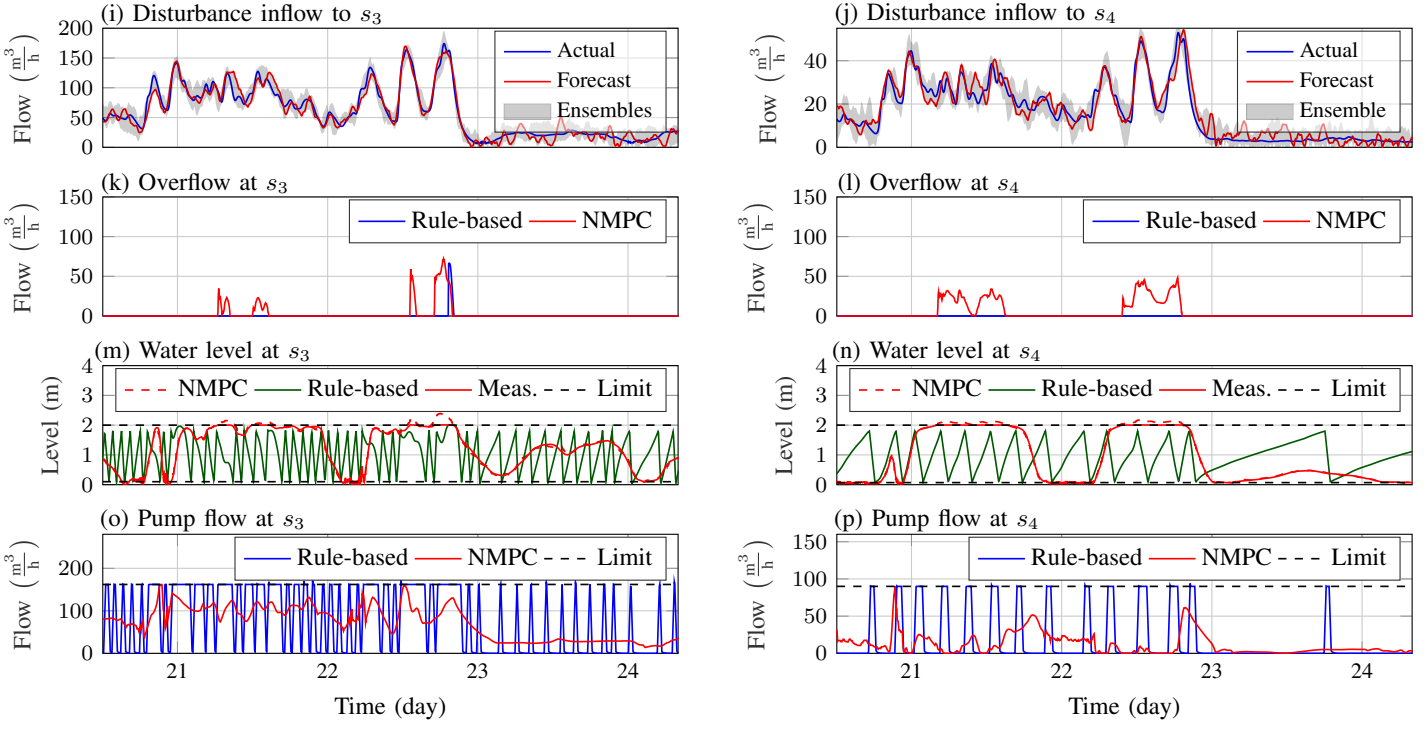


**Figure 9-B:** NMPC with MHPE and MHSE results compared with rule-based control. (a), (b), (i), (j) show disturbance inflows $d_{s,i}$, (c), (d), (k), (l) show overflows $q^{of}_{s,i}$, (e), (f), (m), (n) show states (water levels), while (f), (h), (o), (p) show inputs (pump flows) for each $i \in \{1, ..., N_s\}$ station, respectively.

In the case study, all $\mathbf{\Omega}$ weight values are set equal, hence none of the stations are prioritized over the other. This means that overflow and the filling sensitivities are not prioritized. As the overflows are not avoidable over the selected two days period, the overall goal is to reduce the amount of flooded volume.

The disturbance signals used in the HiFi simulator are historic rain and wastewater flows. To evaluate the NMPC/MHE performance under uncertainty, we generated imperfect forecasts for the internal model of the NMPC. To this end, $n = 10$ different disturbance scenarios have been created by adding normally distributed random data on top of the historic flow signals. As shown in Figure 9-A/B(a), (b), (i), and (j), a set of ensemble of forecasts is produced, indicating a range of possible disturbances. The characterization of the uncertainty for each disturbance is given in the Appendix.

To show the deviation between the prediction by the controller and the state measurement retrieved from the HiFi simulator under uncertainty, we indicated the one-step predictions in Figure 9-A/B(e), (f), (m), and (n) with the dashed red line. Note that the upper constraint is violated under overflow events, due to the slack approximating the volume of overflows. Furthermore, the lower bounds are violated in case the forecasts indicate higher volumes than expected, ending up in the dry-run of the pumps. The NMPC/MHE strategy overflows the upstream tanks ($s_1$, $s_3$ and $s_4$) at times where the rule-based method avoids overflows. This is depicted in Figure 9-A/B(c), (k) and (l), where all storage nodes are prepared by being emptied before the load increases on the network and therefore the controllers distribute the flooded volumes among the corresponding stations, as shown in Figure 9-A/B(e), (m) and (n). Opposed to the rule-based strategy, the overflows are intentional and coordinated, thereby avoiding the overload on the retention tank $s_2$ during the heavy load period. As the system states (i.e., water levels) show, the overflows are shifted in time as the storage nodes attempt to hold back water until their capacity allows. Note that the precise flow-level translation and the precise discharged flow predictions ($y_1$, $y_3$ and $y_4$) guarantee the proper management of the pits ($s_1$, $s_3$, $s_3$) and the retention pit ($s_2$), mitigating the overflow volumes optimally. The comparison of overflow reduction between the baseline and NMPC controllers is shown in Figure 10. Applying the proposed NMPC/MHE strategy results in 28 % cumulative overflow volume decrease over the considered period.

To comprehensively illustrate the practicability of implementing the NMPC/MHE framework, we report on the computational complexity and the dimensions of the optimization problem. To this end, we reduced the network graph shown in Figure 7, by excluding one and two pumping stations, respectively.

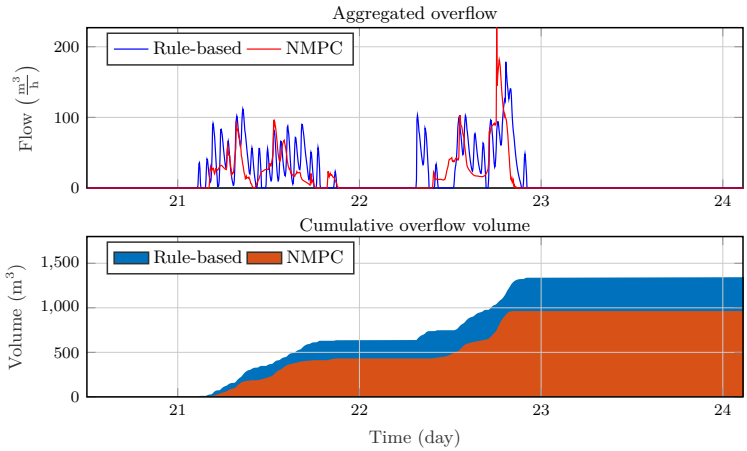


**Figure 10:** Overflow comparison throughout the entire network.

The results, along with the size of the optimization problem, are shown in Table Table 2.

**Table 2:** Computation complexity with different number of stations.

| Num. of stations | Avg. CPU time (s) | Max CPU time (s) | Decision var. | Constr. | Param. |
|---|---|---|---|---|---|
| 4 ($s_{1,2,3,4}$) | 2.14 | 8.37 | 5361 | 8385 | 1216 |
| 3 ($s_{1,2,3}$) | 1.8 | 4.36 | 3912 | 6072 | 912 |
| 2 ($s_{1,2}$) | 1.15 | 1.86 | 2608 | 4192 | 610 |

The optimization problem is carried out on the case study network scaling from two pumping stations to the full extent of the network. As shown, the size of the optimization problem is increasing with including more pumping stations and transport links, however, the computation remains low, as all constraints can be cast as linear equalities and inequalities. Moreover, the average and maximum CPU times for the full scale of the network are only $2.14s$ and $8.37s$. This is acceptable in practice, considering that the worst-case calculations (occurring under overflow events) utilize less than 2 % of the sampling interval $T_{\text{NMPC}} = 10$ (min).

The numerical results carried out on the HiFi network show the feasibility of the proposed data-driven design and provide a basis for onward development. A key outcome of the system identification and control results is that the reduced physically-based SV-PDE flow model can be obtained based on water level measurements, moreover, the discharge predictions are accurately computed via the moving horizon parameter and state estimation.

# 7 Discussion

Our framework aims to allow operators at wastewater utilities to build internal models of the main transport lines and storage nodes in UDNs based on easy-accessible level measurements. Identifying the internal model parameters automatically from standard measurements is therefore one of our contributions. The proposed NMPC/MHE strategy has comparable performance as standard predictive control strategies reported in the literature, benchmarked with rule-based controllers. For instance, references [Joseph-Duran et al., 2014], [Joseph-Duran et al., 2015] report on a hybrid strategy where the internal MPC models exploit all available knowledge from the HiFi network of the UDN. As opposed to [Joseph-Duran et al., 2014], we report on the modularity of our approach, focusing on an internal model obtained by water level data.

Practical implementation of using the method includes the fact that water level sensors need to be deployed in the network to identify the transport dynamics between stations and the periodic household disturbances. Furthermore, our identification approach exploits knowledge about the high-level layout of the network, which is typically available at water utilities. To carry out the experimental implementation of the work, a reliable mapping between rain intensity and the actual flow appearing in the system is needed. Besides, an implementation of a communication strategy is required, where the calculated flow references are translated to reference signals at the local pumping units.

# 8 Conclusions and future work

In this article, a new methodology for data-driven predictive control in urban drainage networks has been presented and tested. The proposed data-driven modeling approach is based on the physical characteristics of open-channel unpressurized flow, governed by the reduced Saint-Venant partial differential equations. A modified version of this model has been used for predicting the internal water levels in the sewer network, moreover to predict the discharged flows to the storage units. To update the model from data, level sensors have been distributed in manholes to enhance the internal prediction performance by taking into account periodic and non-periodic lateral inflows along the pipelines. Moving horizon parameter estimation has been proposed to overcome the inaccuracy issues, introduced by the linearization of the pipe geometries and the approximation of the reduced Saint-Venant partial differential equations. To overcome the problem of limited sensor measurements in the network, moving horizon state estimation has been proposed. The nominal nonlinear multi-objective optimization problem has been solved in a receding horizon fashion, along with the proposed state and parameter estimations. The performance of the proposed methodology has been successfully tested on a high-fidelity sewer network simulator with real rain and domestic wastewater inflow measurement data.

In future work, the methodology will be tested on urban drainage networks with different sizes and topologies. Moreover, it will be interesting to investigate the proposed method in different applications, e.g., stormwater collection networks. Also, an investigation into how rain and domestic wastewater uncertainties can be integrated with the current modeling and control methodology is a matter of future work.

# Appendix

## Simulation parameters

In this appendix we provide the numerical values of the control parameters, the constraint bounds and the main physical attributes of the HiFi simulation network, given in Table 3, Table 4 and Table 5, respectively.

| Attributes | Pipe $g_1$ | Pipe $g_2$ | Pipe $g_3$ | Pipe $g_4$ | Unit |
|---|---|---|---|---|---|
| Geometry | circular | circular | circular | circular | (-) |
| Diameter ($d$) | 0.4 | 0.6 | 0.45 | 0.35 | (m) |
| Slope ($S_b$) | 0.03 | 0.05 | 0.02 | 0.02 | (-) |
| Roughness ($n$) | 0.013 | 0.013 | 0.013 | 0.013 | (-) |
| Length ($L$) | 0.9 | 0.45 | 2 | 2.4 | (km) |

**Table 3:** Network attributes for pipes.

Note that the values for the upper height constraints $\overline{h}$ are equivalent to the diameter of the pipes.

| Attributes | Pit $s_1$ | Pit $s_2$ | Pit $s_3$ | Pit $s_4$ | Unit |
|---|---|---|---|---|---|
| Constant ($K_1$) | 21.5 | 30 | 30 | 43 | ($m^2$) |
| Constant ($K_2$) | - | 130 | - | - | ($m^2$) |
| Volume ($\overline{V}$) | 43 | 95 | 60 | 86 | ($m^3$) |
| Pump flow ($\overline{u}$) | 162 | 198 | 162 | 90 | ($m^3$/h) |

**Table 4:** Network attributes for pits.

| $T_{\text{NMPC}}$ | $T_{\text{on/off}}$ | $H_{p1}$ | $H_{p2}$ | $H_{pe}$ | $H_{se}$ | $T_{se}$ |
|---|---|---|---|---|---|---|
| 10 (m) | 1 (m) | 2 (h) | 22 (h) | 48 (h) | 48 (h) | 10 (m) |

**Table 5:** Simulation parameters.

The disturbance model uses $k = 2$ frequency terms and $\omega = 1$ (day) frequency for all pipes. Moreover, the disturbance signal scenarios are characterized by normally distributed, zero mean random uncertainty, where $\sigma^2_{s_1} = 30.6$, $\sigma^2_{s_2} = 5.4$, $\sigma^2_{s_3} = 27$ and $\sigma^2_{s_4} = 10.8$ $(m^3/h)$. The lateral inflows along the gravity pipelines are all characterized by $\sigma^2_g = 15$ $(m^3/h)$.

## Acknowledgment

The authors would like to thank Rasmus Halvgaard from DHI Group for helping with setting up the MIKE 1D API to the MIKE URBAN simulation environment. The work of C. Ocampo-Martinez has been supported by the project PID2020-115905RB-C21 (L-BEST) funded by MCIN/ AEI /10.13039/50110001 1033. Furthermore, we thank the wastewater utility in Fredericia, Denmark for providing real flow data for our numerical experiment.

## References

F. Allgöwer, T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright. "Nonlinear Predictive Control and Moving Horizon Estimation — An Introductory Overview". In *Advances in Control*, pages 391–449. Springer, London, 1999. doi:10.1007/978-1-4471-0853-5_19.

J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. "CasADi: a software framework for nonlinear optimization and optimal control". *Mathematical Programming Computation*, 11(1):1–36, 2019. ISSN 18672957. doi:10.1007/s12532-018-0139-4.

K. M. Balla, C. S. Kallesoe, C. Schou, and J. D. Bendtsen. "Nonlinear grey-box identification with inflow decoupling in gravity sewers". *IFAC-PapersOnLine*, 53 (2):1065–1070, 2020a. ISSN 24058963. doi:10.1016/j.ifacol.2020.12.1295.

K. M. Balla, C. Schou, J. Dimon Bendtsen, and C. S. Kallesoe. "Multi-scenario model predictive control of combined sewer overflows in urban drainage networks". In *2020 IEEE Conference on Control Technology and Applications (CCTA)*, pages 1042–1047, Montréal, 2020b. IEEE. ISBN 9781728171401. doi:10.1109/CCTA41146.2020.9206362.

D. Butler and J. W. Davies. *Urban Drainage*. Spon Press, 2006. doi:10.1016/s1462-0758(00)00017-0.

F. J. Chang, J. M. Liang, and Y. C. Chen. "Flood forecasting using radial basis function neural networks". *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 31(4):530–535, 2001. ISSN 10946977. doi:10.1109/5326.983936.

V. Dalmas, G. Robert, G. Besançon, and D. Georges. "Simplified Non-Uniform Models for Various Flow Configurations in Open Channels". *IFAC-PapersOnLine*, 50(1):12320–12325, 2017. ISSN 24058963. doi:10.1016/j.ifacol.2017.08.2159.

S. Dey. "Free overfall in open channels: State-of-the-art review". *Flow Measurement and Instrumentation*, 13(5-6):247–264, 2002. ISSN 09555986. doi:10.1016/S0955-5986(02)00055-9.

L. García, J. Barreiro-Gomez, E. Escobar, D. Téllez, N. Quijano, and C. Ocampo-Martinez. "Modeling and real-time control of urban drainage systems: A

review". *Advances in Water Resources*, 85:120–132, 2015. ISSN 03091708. doi:10.1016/j.advwatres.2015.08.007.

M. S. Gelormino and N. L. Ricker. "Model-predictive control of a combined sewer system". *International Journal of Control*, 59(3):793–816, 1994. ISSN 13665820. doi:10.1080/00207179408923105.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. "Hybrid modeling and receding horizon control of sewer networks". *Water Resources Research*, 50(11): 8497 – 8514, 2014. ISSN 19447973. doi:10.1002/2013WR015119.

B. Joseph-Duran, C. Ocampo-Martinez, and G. Cembrano. "Output-feedback control of combined sewer networks through receding horizon control with moving horizon estimation". *Water Resources Research*, 51(10):8129–8145, 2015. ISSN 19447973. doi:10.1002/2014WR016696.

C. S. Kallesøe and T. Knudsen. "Self calibrating flow estimation in waste water pumping stations". In *2016 European Control Conference, ECC 2016*, pages 55–60, 2016. ISBN 9781509025916. doi:10.1109/ECC.2016.7810263.

X. Litrico and V. Fromion. "Boundary control of linearized Saint-Venant equations oscillating modes". *Automatica*, 42:967–972, 2006. ISSN 00051098. doi:10.1016/j.automatica.2006.02.002.

A. M. D. Livera, R. J. Hyndman, and R. D. Snyder. "Forecasting time series with complex seasonal patterns using exponential smoothing". *Journal of the American Statistical Association*, 106(496):1513–1527, 2011. doi:10.1198/jasa.2011.tm09771.

R. Löwe, S. Thorndahl, P. S. Mikkelsen, M. R. Rasmussen, and H. Madsen. "Probabilistic online runoff forecasting for urban catchments using inputs from rain gauges as well as statically and dynamically adjusted weather radar". *Journal of Hydrology*, 512:397–407, 2014. ISSN 00221694. doi:10.1016/j.jhydrol.2014.03.027.

R. Löwe, L. Vezzaro, P. S. Mikkelsen, M. Grum, and H. Madsen. "Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems". *Environmental Modelling and Software*, 80:143–158, 2016. ISSN 13648152. doi:10.1016/j.envsoft.2016.02.027.

N. S. V. Lund, A. K. V. Falk, M. Borup, H. Madsen, and P. Steen Mikkelsen. "Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management". *Critical Reviews in Environmental Science and Technology*, 48(3):279–339, 2018. ISSN 15476537. doi:10.1080/10643389.2018.1455484.

J. Ma, W. Sun, G. Yang, and D. Zhang. "Hydrological Analysis Using Satellite Remote Sensing Big Data and CREST Model". *IEEE Open Access Journals and Conferences*, 6:9006–9016, 2018. ISSN 21693536. doi:10.1109/ACCESS.2018.2810252.

L. W. Mays. *Stormwater Collection Systems Design Handbook*. McGraw-Hill Education, 1st edition, 2001. ISBN 0071354719.

MIKE powered by DHI. "MIKE for developers". Technical report. URL: `https://docs.mikepoweredbydhi.com/engine_libraries/mike1d/mike1d_api/`.

MIKE powered by DHI. "Collection System - Modelling of storm water drainage networks and sewer collection systems". Technical report, Hørsholm, 2017.

MIKE powered by DHI. "MIKE 1D: DHI Simulation Engine for 1D river and urban modelling - Reference Manual". Technical report, DHI, 2019.

A. L. Mollerup, P. S. Mikkelsen, and G. Sin. "A methodological approach to the design of optimising control strategies for sewer systems". *Environmental Modelling and Software*, 83:103–115, 2016. ISSN 13648152. doi:10.1016/j.envsoft.2016.05.004.

S. Munier, X. Litrico, G. Belaud, and P. O. Malaterre. "Distributed approximation of open-channel flow routing accounting for backwater effects". *Advances in Water Resources*, 31(12):1590–1602, 2008. ISSN 03091708. doi:10.1016/j.advwatres.2008.07.007.

C. Ocampo-Martinez. *Model Predictive Control of Wastewater Systems*. Springer, Barcelona, 1st edition, 2010. ISBN 978-1-84996-352-7. doi:10.1007/978-1-84996-353-4.

C. Ocampo-Martinez, V. Puig, G. Cembrano, and J. Quevedo. "Application of predictive control strategies to the management of complex networks in the urban water cycle". *IEEE Control Systems*, 33(1):15–41, 2013. ISSN 1941000X. doi:10.1109/MCS.2012.2225919.

J. A. Roberson and C. T. Crowe. *Engineering Fluid Mechanics*. Houghton Mifflin Company, Boston, 5th edition, 1993. ISBN 0-395-66161-7. doi:10.1007/978-1-4020-6742-6.

M. R. Schütze, D. Butler, and M. B. Beck. *Modelling, Simulation and Control of Urban Wastewater Systems*. Springer, 2002. doi:10.1007/978-1-4471-0157-4.

X. Tian, R. R. Negenborn, P. J. van Overloop, J. María Maestre, A. Sadowska, and N. van de Giesen. "Efficient multi-scenario Model Predictive Control for water resources management with ensemble streamflow forecasts". *Advances in Water Resources*, 109:58–68, 2017. ISSN 03091708. doi:10.1016/j.advwatres.2017.08.015.

S. C. Troutman, N. Schambach, N. G. Love, and B. Kerkez. "An automated toolchain for the data-driven and dynamical modeling of combined sewer systems". *Water Research*, 126:88–100, 2017. ISSN 18792448. doi:10.1016/j.watres.2017.08.065.

K.-J. van Heeringen, R. van Nooijen, K. Kooij, and B. Postma. "Real-time Control of sewer pumps by using ControlNEXT to smooth inflow at Waste Water Treatment Plant Garmerwolde". In *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, pages EPSC2016–1412, Apr. 2016.

A. Wächter and L. T. Biegler. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming". *Mathematical Programming*, 106(1):25–57, 2006. ISSN 00255610. doi:10.1007/s10107-004-0559-y.

A. Wills and B. Ninness. "On gradient-based search for multivariable system estimates". *IEEE Transactions on Automatic Control*, 53(1):298–306, 2008. ISSN 00189286. doi:10.1109/TAC.2007.914953.

M. Xu, P. J. van Overloop, and N. C. van de Giesen. "On the study of control effectiveness and computational efficiency of reduced Saint-Venant model in model predictive control of open channel flow". *Advances in Water Resources*, 34:282–290, 2011. ISSN 03091708. doi:10.1016/j.advwatres.2010.11.009.

M. Xu, R. R. Negenborn, P. J. van Overloop, and N. C. van de Giesen. "De Saint-Venant equations-based model assessment in model predictive control of open channel flow". *Advances in Water Resources*, 49:37–45, 2012. ISSN 03091708. doi:10.1016/j.advwatres.2012.07.004.

Y. Zou, L. Cen, D. Li, and X. He. "Simplified state-space model and validation of irrigation canal systems". In *Chinese Control Conference, CCC*, pages 2002–2007, Hangzhou, 2015. ISBN 9789881563897. doi:10.1109/ChiCC.2015.7259938.